DOCUMENT RESUME

ED 144 398 FL 008 902

AUTHOR Flahive, Douglas E.
TITLE Separating the (g) Factor from Reading Comprehension. Occasional Papers on Linguistics, No. 1.
INSTITUTION Southern Illinois Univ., Carbondale. Dept. of Linguistics.
PUB DATE Apr 77
NOTE 7p.; Proceedings of the International Conference on Frontiers in Language Proficiency and Dominance Testing (1st, Southern Illinois University, Carbondale, Illinois, April 21-23, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Cloze Procedure; *English (Second Language); *Intelligence; Intelligence Tests; Language Instruction; Language Proficiency; Language Skills; *Language Tests; Multiple Choice Tests; Predictive Ability (Testing); *Reading Comprehension; Reading Skills; *Reading Tests; Second Language Learning; *Test Validity
IDENTIFIERS *Language Comprehension

ABSTRACT

This paper discusses the role of intelligence in language comprehension. In the study reported, an attempt is made to determine how much of the variability in English reading scores of non-native speakers of English can be predicted on the basis of intelligence alone. In the study, three measures of reading comprehension are used: a traditional multiple-choice test, in which a reading passage is followed by a series of questions; a paraphrase recognition test, which requires that the testee choose from among four choices the sentence closest to the probe; and a cloze test, which requires the testee to fill in the blanks in a passage with words that are either close to or identical to those omitted. It is predicted that of the three tests, the traditional multiple-choice test will be more highly correlated with intelligence than the other two measures, because traditional multiple-choice requires reasoning ability: the ability to retain facts, to draw conclusions, and to use contextual clues to determine the meaning of unknown words. Results allow the conclusion that reading sub-tests found on widely used measures of non-native speaker language ability are tests which, to some degree, measure intelligence, and the validity of the tests is therefore called into question. (Author/CLK)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED144398

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Southern Illinois University
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

R008902

Separating the (g) Factor from Reading Comprehension

Douglas E. Flahive

Introduction

This paper addresses topics that have been focal points of extensive educational research since the beginning of this century: intelligence and language comprehension. Despite this extensive research, little agreement exists among researchers concerning even the basic definitions of the terms. Is intelligence a static construct measured by performance on standardized I.Q. tests as Jensen claims, or is it a dynamic construct as Piaget and his followers claim? And what about language comprehension? What is it, and how is it measured? If one reads and attempts to synthesize the numerous studies that have addressed tne second question, the frustrating conclusion would be that there is no universally accepted definition of language comprehension, nor is there a universally valid technique for assessing language comprehension.

While attempts to define and measure language comprehension have many numerous and varied, a persistent question runs through many of the efforts. The question concerns the role of intelligence in language comprehension. Based on the pioneering work of Davis(1944), reading comprehension skills rely greatly on skills associated with intelligence. Carroll(1971) in summarizing nearly three decades of research into questions of language comprehension concludes that reading comprehension is dependent on two factors-grammatical knowledge and inferential ability.

In the study reported below, an attempt is made to determine how much of the variability in English reading scores of non-native speakers of English can be predicted on the basis of intelligence alone. In this study, three measures of reading comprehension are used: a traditional multiple-choice test in which a reading passage is followed by a series of questions, a paraphrase recognition test which requires that the testee choose from among four choices the sentence which is the closest to the probe, and a cloze test which requires the testee to fill in blanks of a passage with words that either are close to or identical to those omitted. It is predicted that of the three

tests of reading comprehension, the traditional multiple-choice test will be more highly correlated with intelligence than the other two measures. This is because the traditional multiple-choice requires the testee to do much more than simply comprehend language. It requires the ability to retain facts, to draw conclusions, to use contextual clues to determine the meaning of unknown words. In short, it requires reasoning ability.

Subjects and Materials

Twenty students enrolled in a semi-intensive English class were administered five tests over a one week period. The students represent seven different native language backgrounds. The first test they were administered was the TOEFL(Test of English as a Foreign Language). The purpose of this test is to measure the overall English proficiency of non-native speakers. The next test administered was the Raven's Progressive Matrices(Standard Form). This test was selected for this study for two major reasons. First, it has proved to be a reliable measure of intelligence in a wide variety of non-English speaking settings. [1] Second, the test is one of the few language free intelligence tests. Using a language free test of intelligence seemed to be a far more practical alternative to locating and administering comparable intelligence tests in the native language of the students. The test requires that the testee look at a pattern which has a portion of it missing and select from among the eight choices below the pattern which would best complete the larger pattern. The test contains five sections with twelve items in each section. The items get progressively more difficult as the subject advances through each section. As an added control to make sure that the test results were free from language bias and that the subjects fully understood the task requirements, directions were given in the native language of the subjects. The test was group-administered, and the subjects given 45 minutes to complete it. A possible maximum score is 60.

In addition to the Raven's and the TOEFL, three reading comprehension tests were administered: the paragraph comprehension portion of the McGraw-Hill Basic Skills System Reading Test developed by A.L. Raygor(1970), a paraphrase recognition test developed by K.Perkins and C. Yorio(See Perkins(1976) for details), and a cloze test developed by the author. The McGraw-Hill Test is designed to measure a subject's ability to make inferences, pick out main and supporting ideas from a paragraph, and

discover organizational patterns of paragraphs and essays. The test contains a total of 10 passages-5 long(approximately 250 words for each) and 5 short(approximately 75 words for each) followed by a total of 30 multiple-choice questions. The subjects were given 40 minutes to complete the test. The Perkins - Yorio test is a 50 item paraphrase identification test with items designed to measure comprehension of specific grammatical structures. A sample test item is illustrated below.

Flying airplanes can be dangerous.
A. Airplanes that are flying can be dangerous.
B. Airplanes that are flying are safe.
C. To fly airplanes is safe.
D. To fly airplanes is not dangerous at all.

Subjects were given 30 minutes to complete the 50 items. The cloze test is a 400 word passage taken from an introductory economics text and slightly modified by the author. After several lead sentences, every 7th word was deleted for a total of 50 possible correct answers. The appropriate word scoring method was used. The subjects were given 30 minutes to complete the test.

Results

The means and standard deviations for each of the five tests are presented in Table 1 below. The score obtained on

Table 1
Means and standard deviations for the 20 subjects over the 5 tests.

Raven's	Toefl	M-H Test	Perkins-Yorio	Cloze
49.90	509.40	15.50	41.20	18.85
6.34	38.80	4.28	4.63	7.44

Raven's would place them slightly above the 75th percentile of the group who formed the reference group for the test (Raven, 1960). One can reasonably conclude that the subjects were above average intelligence. The reader is free to interpret the degree of English language proficiency which

is represented by the TOEFL score. Given the range of scores, 568-437, the only firm conclusion that can be drawn is that the subjects ranged from highly to marginally proficient in overall English skills. The McGraw-Hill Test scores would place students from the 85th to the 9th percentile when compared to a reference group of native speaking college freshmen and sophmores with the mean of the group at the 22nd percentile. Performance on the Perkins-Yorio test was only slightly below scores of native speaking college freshmen.

The correlations among the five measures are seen in Table 2 below. The fact that all the measures are

Table 2
Correlation matrix for the five measures.

	Raven's 1	TOEFL 2	Perkins-Yorio 3	Cloze 4	McGraw-Hill 5
1.	1.0				
2.	.61	1.0			
3.	.68	.84	1.0		
4.	.61	.75	.76	1.0	
5.	.84	.59	.67	.62	1.0

p .01 for all correlations

highly correlated is not surprising since three of the tests are reading tests and purportedly are testing the same process. Nor are the correlations between Perkins-Yorio and TOEFL and Cloze and TOEFL surprising. The are all measures of language proficiency. What is surpsising is the high correlation between the Raven's and the McGraw-Hill test. Over 70% of the variance in the McGraw-Hill Test is accounted for by a non-verbal intelligence test. To contrast the predictive power of the Raven's over the three reading tests, three regression analyses were run with the Raven's and TOEFL as predictors in the full models and with the Raven's alone in

the restricted models. The results of these analyses are seen in Table 3. As the reader can see, almost as much of

Table 3

Results of regression analyses using Raven's and TOEFL to predict reading scores for the three measures.

	R^2
Raven's and TOEFL to predict M-H	.716
Raven's alone to predict M-H	.707
Raven's and TOEFL to predict P-Y	.754
Raven's alone to predict P-Y	.457
Raven's and TOEFL to predict Cloze	.598
Raven's alone to predict Cloze	.371

variance accounted for by the Raven's and TOEFL combined on the McGraw-Hill Test was accounted for by the Raven's alone. However in attempting to account for the variance in the Perkins-Yorio and the Cloze tests, significantly less was accounted for by the Raven's alone. From a description of the tests given above, the reason for these results are rather clear. Neither of the two requires the type of complex reasoning involved in attempting to answer the type of questions asked in traditional multiple-choice tests.

Discussion

With the exception of the relatively high ability to account for reading scores by means of a non-verbal test of intelligence, the general results of this study are, for the most part, not unexpected. Based on these results, it is reasonable to conclude that reading sub-tests found on widely used measures of non-native speaker language ability are not simply tests of reading ability. They are tests which, to some degree, measure intelligence. If the multiple-choice format reading comprehension test which is found on both the Michigan and the TOEFL tests is also an intelligence measure, the validity of such measures must be questioned. In a recent study, Johnson(1977) reports that when the TOEFL was administered to native speakers, subscores for listening comprehension, grammar, and vocabulary were close to maximum. For the reading subscore, scores were

significantly lower.From this another other evidence as well, it is clear the traditional tests of reading comprehension are testing not simply language proficiency but intelligence as well(Genesee, 1976).

While it would be rash to suggest that all traditional multiple-choice reading tests are simply disguised intelligence tests, the findings reported in this study suggest that additional studies be done to determine more precisely the role of intelligence in reading comprehension. In addition, the findings also suggest that the role of intelligence in other areas of language processing and in other techniques of language testing be examined to provide teachers and researchers alike with a clearer picture of where language ability or proficiency ends and intelligence begins.

Note

1 The author of the Raven's claims it is not a true test of intelligence but rather a "test of observation and clear thinking". The he goes on to say that it correlates at the 0.86 level with the Terman-Merrill scale and has a "g" saturation of 0.82. The reader is left to conclude whether the Raven's is or is not an intelligence test.